1. AULA 4

Variáveis Aleatórias Contínuas

1.1. **Variáveis Aleatórias Contínuas.** A título de introdução e para observar como os modelos contínuos aparecem naturalmente na prática, recordemos, do capítulo anterior, um exemplo:

Exemplo 1.1. Um engenheiro inspeciona determinada máquina em períodos discretos de tempo e observa se a máquina quebrou durante o período anterior. Independente do período, a probabilidade da máquina quebrar em determinado período é $p, 0 < p < 1$. A probabilidade da máquina não quebrar até o n-ésimo período foi calculada e é igual a $(1-p)^n$.

Suponha, agora, que o engenheiro inspeciona a máquina continuamente no tempo e que as condições para que as falhas ocorram aleatóriamente sejam verdadeiras, isto é, se $N(t)$ conta o número de falhas em $(0,t]$, as suposições I, II, III e IV, das ocorrências aleatórias no tempo, do capítulo anterior, são satisfeitas.

Se dividimos o intervalo $(0,t]$ em intervalos infinitesimais da forma $(\frac{kt}{n}, \frac{(k+1)t}{n}], 0 \leq k \leq n-1$, de comprimento $\frac{t}{n}$, concluimos que, como na aproximação da distribuição binomial pela distribuição de Poisson demonstrada no Capítulo anterior, $N(t)$, o número de ocorrências em $(0,t]$, tem distribuição binomial de parâmetros n e $p = \frac{\lambda.t}{n} + o(\frac{t}{n})$.

Seja, agora, T o tempo da ocorrência do primeiro evento de $N(t)$. Claramente, quando $n \to \infty$,

$$P(T > t) = P(N(t) = 0) = (1 - \frac{\lambda.t}{n} - o(\frac{t}{n}))^n \to \exp[-\lambda.t].$$

Portanto, a função de distribuição de T é $F_T(t) = 1 - \exp[-\lambda.t]$, da distribuição exponencial de parâmetro λ, definida em $\Re^+$, é contínua e que estudaremos, com mais detalhes, nas próximas seções.

Passemos à definição da variável aleatória contínua com argumentos semelhantes àqueles da definição das variáveis aleatórias discretas. Contudo, se o espaço amostral Ω é o resultado de um experimento quantitativo contínuo, a imagem de Ω através de uma função real X não é, certamente, enumerável e não podemos definir $\Im_X$ como o conjunto das partes de $X(\Omega)$.

Para termos uma noção de como abordar a questão, observe que qualquer subconjunto dos números reais pode ser obtido através de operações, em um número finito ou infinito enumerável, de intervalos da forma $(-\infty, t]$. Entre outras operações, exemplificamos:

$(t, \infty) = \overline{(-\infty, t]}$;
$(s, t] = (-\infty, t] - (-\infty, s], s < t$;
$\{t\} = \cap_{n=1}^{\infty}(t - \frac{1}{n}, t + \frac{1}{n}]$.

Definição 1.2. Definimos β como sendo a classe de subconjuntos dos reais obtida através das operações de reunião, intersecção, complementar, em número finito ou infinito, de subconjuntos na forma $(-\infty, t], \forall t \in \Re$. β é denominada σ álgebra de Borel na reta.

Nesta classe definimos a medida de probabilidade induzida pela variável aleatória $X : \Omega \to \Re$. Definimos a probabilidade induzida por X, P_X, como $P_X(B) = P(X^{-1}(B)), \forall B \in \beta$, **quando** X satisfaz a seguinte definição:

Definição 1.3. Seja $(\Omega, \Im, P)$ um espaço de probabilidade e X uma aplicação de Ω nos reais ($\Re$). X é uma variável aleatória contínua se $X^{-1}((0, t]) \in \Im, \quad \forall t \in \Re$. Denominamos $(\Re, \beta, P_X)$ como o espaço de probabilidade induzido por X.

Observe que P_X esta bem definida e

$$P_X((-\infty, t]) = P(X^{-1}((-\infty, t])) = P(\{w : X(w) \in (-\infty, t]) =$$
$$P(X \leq t) = F_X(t)$$

é a função de distribuição da variável aleatória X que , neste caso, é uma função contínua.

Se a função de distribuição de X é diferenciável com $\frac{dF_X(t)}{dt} = f_X(t)$, definimos

Definição 1.4. Se X é uma variável aleatória contínua com função de distribuição

$$F_X(t) = \int_{-\infty}^{t} f_X(y)dy,$$

dizemos que $f_X(t)$ é a função densidade de probabilidade de X e que X é contínua. A sua função de distribuição é denominada absolutamente contínua.

Observação 1.5. O valor $F_X(t) = \int_{-\infty}^{t} f_X(y)dy$ não muda se $f_X(y)$ apresentar um conjunto enumerável de ponto de descontinuidade de maneira que, diferente do caso discreto, $F(x)$ não determina completamente a função densidade de probabilidade.

Observe que

$$f(x) = \frac{dF(x)}{dx} = \lim_{\triangle x \to 0} \frac{F(x + \triangle x) - F(x - \triangle x)}{2\triangle x}$$

e portanto
$$2\triangle x f(x) \approx P(x - \triangle x < X \leq x + \triangle x),$$
a probabilidade de que X está em um intervalo infinitesimal contendo o valor x é igual a $f(x)$ vezes o seu comprimento, isto é, a área do retângulo de base de comprimento infinitesimal $2\triangle x$ e altura $f(x)$.

É evidente que a função densidade de probabilidade é positiva, isto é, $f_X(t) \geq 0, \forall t$, que
$$\int_{-\infty}^{\infty} f(x)dx = 1$$
e que
$$P(a < X \leq b) = \int_a^b f(x)dx.$$

Na segunda aula, provamos que se X é uma variável aleatória, então
$$\forall \varepsilon > 0. \exists k \in \aleph : P(|X| > k) < \varepsilon.$$
e, usando este argumento, podemos provar que
$$\lim_{t \downarrow -\infty} f_X(t) = \lim_{t \uparrow \infty} f_X(t) = 0.$$

Exemplo 1.6. O custo mensal dos sinistros de uma Uma Cia. de Seguros é modelado por uma variável aleatória X com função densidade de probabilidade
$$f(x) = \begin{cases} 0 & : \quad x < 0 \\ k.(1+x)^{-4} & : \quad 0 < x < \infty \end{cases},$$
onde k é uma constante.

Como devemos ter $\int_{-\infty}^{\infty} f(x)dx = 1$ e
$$1 = k \int_0^{\infty} (1+x)^{-4} dx = -\frac{k}{3}(1+x)^{-3}|_0^{\infty} = \frac{k}{3},$$
temos $k = 3$.

A função de distribuição de X é
$$F(x) = \begin{cases} 0 & : \quad x < 0 \\ \int_0^x 3(1+y)^{-4} dy = 1 - (1+x)^{-3} & : \quad x \geq 0 \end{cases}.$$

A probabilidade condicional de que o custo mensal ultrapasse 40 conhecendo-se que é maior do que 10 é
$$P(X > 40 | X > 10) = \frac{P(X > 40, X > 10)}{P(X > 10)} = \frac{P(X > 40)}{P(X > 10)} =$$

$$\frac{3\int_{40}^{\infty}(1+y)^{-4}dy}{3\int_{10}^{\infty}3(1+y)^{-4}dy} = (\frac{11}{41})^3 = 0,02.$$

Lembrando que as integrais de Riemann são limites de somas infinitesimais é natural que os parâmetros de uma variável aleatória contínua sejam definidos de maneira semelhante aos da variável aleatória discreta:

Definição 1.7. Se X é uma variável aleatória contínua com função densidade de probabilidade $f(x)$, definimos a média de X, que denotamos por $E[X]$ ou μ, à soma

$$\mu = \int_{-\infty}^{\infty} xf(x)dx,$$

onde a integral deve ser absolutamente convergente, caso contrário dizemos que a média não existe.

Por vezes estamos interessados em variáveis aleatórias que resultam da transformação de outra variável. Por exemplo, seja X uma variável aleatória e $Y = g(X)$, onde g é uma função real. Pense no esquema

$$(\Omega, \Im, P) \Rightarrow (\Re, \Im_X, P_X) \Rightarrow (\Re, \Im_{g(X)}, P_{g(X)}).$$

A medida $P_{g(X)}$ é caracterizada por

$$P_{g(X)}((-\infty, y]) = P(g(X) \leq y) = \int_{\{x:g(x)\leq y\}} f_X(x)dx.$$

Pode-se provar que

$$E[g(X)] = E[Y] = \int_{-\infty}^{\infty} yf_Y(y)dy = \int_{-\infty}^{\infty} g(x)f_X(x)dx.$$

Exemplo 1.8. No exemplo anterior, em que

$$f(x) = \begin{cases} 0 & : \quad x < 0 \\ 3.(1+x)^{-4} & : \quad 0 < x < \infty \end{cases},$$

$$\mu = E[X] = 3\int_0^{\infty} x(1+x)^{-4}dx = 3\int_1^{\infty}(y-1)y^{-4}dy = 3.(\frac{1}{2}-\frac{1}{3}) = \frac{1}{2}.$$

$$E[X^2] = 3\int_0^{\infty} x^2(1+x)^{-4}dx = 2\int_0^{\infty} x(1+x)^{-3}dx = \frac{1}{3}\int_0^{\infty}(1+x)^{-2}dx = 1.$$

Em particular a esperança da n-ésima potência de variável aleatória, se existir, é denominada momento.

Definição 1.9. Seja X uma variável aleatória contínua. O n-ésimo momento de X, denotado por μ_n é definida como:

$$\mu_n = E[X^n] = \int_{-\infty}^{\infty} x^n f(x)dx, \quad se \ \ existir.$$

Observe que $\mu_1 = \mu = E[X]$. A esperança de potências da forma $E[(X-\mu)^n]$ são denominadas momentos centrais.

Definição 1.10. Seja X uma variável aleatória contínua com média μ. O n-ésimo momento central de X, denotado por σ_n é definida como:

$$\sigma_n = E[(X-\mu)^n)] = \int_{-\infty}^{\infty} (x-\mu)^n f(x)dx, \quad se \ \ existir.$$

Em particular o segundo momento central em torno da média é a variância, interpretada como uma dispersão em torno da média.

Definição 1.11. Se X é uma variável aleatória contínua com média μ e função de densidade de probabilidade $f(x)$, definimos a variância de X, que denotamos por $E[(X-\mu)^2] = \sigma^2$, à integral

$$\sigma_2 = \sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx.$$

relembrando que a integral deve ser absolutamente convergente, caso contrário dizemos que a variância não existe.

Observação 1.12. Como no caso discreto, é facil provar que $\sigma^2 = Var(X) = E[X^2] - \mu^2$. O desvio padrão de X é definido pela raiz quadrada de σ^2. No exemplo anterior temos que $Var(X) = 1 - \frac{1}{4} = \frac{3}{4}$ e $dP(X) = \frac{\sqrt{3}}{2}$.

Definição 1.13. Se X é uma variável aleatória contínua . A função geradora de momentos de X, $M_X(t)$, é o valor esperado $E[\exp[tX]]$, se existir, em um intervalo simétrico, $(-s,s)$, de números reais.

$$M_X(t) = E[\exp[tX]] = \int_{-\infty}^{\infty} \exp[tx]f(x)dx.$$

Exemplo 1.14. Em uma Cia. de manufaturas, as perdas por danos à propriedade segue uma distribuição contínua com função densidade de probabilidade $f(y) = 0,02\exp[-0,02y], y > 0$. A função geradora de momentos de Y é

$$M_Y(t) = E[\exp[tY]] = \int_{0}^{\infty} \exp[ty]0,02\exp[-0,02y]dy =$$

$$\int_0^{\infty} 0,02\exp[(t-0,02)y]dy = \frac{0,02}{0,02-t}$$

, se $t < 0,02$.

A primeira derivada de $M_Y(t)$ é $M_Y^{(1)}(t) = \frac{0,02}{(0,02-t)^2}$ e

$$E[Y] = M_Y^{(1)}(0) = \frac{1}{0,02}.$$

A segunda derivada de $M_Y(t)$ é $M_Y^{(2)}(t) = \frac{2.(0,02)}{(0,02-t)^3}$ que no número 0 vale

$$E[Y^2] = M_Y^{(2)}(0) = \frac{2}{(0,02)^2}.$$

Portanto a variãncia de Y é $\sigma^2 = E[Y^2] - E[Y]^2 = \frac{1}{(0,02)^2}$.

Observação 1.15. As desigualdades de Markov, Chebyshev e de Jensen continuam valendo no caso de variáveis aleatórias contínuas e as provas são semelhantes.

1.2. **Modelos Probabilísticos Contínuos.** Como no capítulo anterior, descreveremos modêlos probabilísticos contínuos aplicáveis nas ciencias atuariais:

Distribuição uniforme.

Em um espaço amostral discreto e finito, a distribuição uniforme associa a cada elemento amostral a mesma probabilidade. Em um modelo contínuo cada ponto tem probabilidade igual a zero e a distribuição uniforme se caracteriza associando a cada intervalo de mesmo comprimento a mesma probabilidade. A variável aleatória X tem distribuição uniforme em um intervalo finito, $(a, b]$, de números reais se sua função densidade de probabilidade é :

$$f(x) = \begin{cases} \frac{1}{b-a} & : \quad a < x \le b \\ 0 & : \quad c.c. \end{cases}.$$

A esperança de X é

$$\mu = E[X] = \int_a^b \frac{x}{b-a}dx = \frac{1}{b-a}\frac{x^2}{2}|_a^b = \frac{(b-a).(b+a)}{2(b-a)} = \frac{a+b}{2}.$$

Tambem temos

$$E[X^2] = \int_a^b \frac{x^2}{b-a}dx = \frac{1}{b-a}\frac{x^3}{3}|_a^b = \frac{(b-a).(b^2+ab+a^2)}{3(b-a)} = \frac{b^2+ab+a^2}{3}$$

e portanto

$$Var(X) = \sigma^2 = E[X^2] - \mu^2 = \frac{b^2 + ab + a^2}{3} - (\frac{a+b}{2})^2 = \frac{(b-a)^2}{12}.$$

A função de distribuição de X é

$$F(x) = \begin{cases} 0 & : \quad x < a \\ \frac{x-a}{b-a} & : \quad a \leq x < b \\ 1 & : \quad x \geq b \end{cases}.$$

e a função geradora de momentos

$$M_X(t) = E[\exp[tX]] = \int_a^b \frac{\exp[tX]}{b-a} dx = \frac{\exp[tb] - \exp[ta]}{t(b-a)}.$$

Exemplo 1.16. Uma seguradora de automóveis cobra $R\$250,00$ pela franquia de determinada apólice e paga o máximo de $R\$1.500,00$ por uma perda total. Se o custo dos sinistros, em relação à apólice, é modelado por uma distribuição uniforme no intervalo $(0, 2.000]$, qual a probabilidade de um pagamento de perda total? Qual a probabilidade do segurado não usar a apólice?

Se X é a variável custo, tem distribuição uniforme no intervalo $(0, 2000]$, com função densidade de probabilidade

$$f(x) = \begin{cases} \frac{1}{2.000} & : \quad 0 < x \leq 2.000 \\ 0 & : \quad c.c. \end{cases}.$$

A probabilidade de um pagamento de perda total é

$$P(X > 1.500) = \int_{1.500}^{2000} \frac{1}{2.000} dx = 0,25.$$

Entendemos que o segurado não usa a apólice quando a franquia é maior do que o custo do conserto e portanto, com probabilidade

$$P(X \leq 250) = \int_0^{250} \frac{1}{2.000} dx = 0,125.$$

Distribuição de Cauchy.

A variável aleatória de Cauchy é caracterizada pela função densidade de probabilidade

$$f(x) = \frac{1}{\pi\beta\{1 + [(\frac{x-\alpha}{\beta})]^2\}}, \quad -\infty < x < \infty,$$

com parâmetros α, $\; -\infty < \alpha < \infty$ e β, $\; \beta > 0$.

Em particular se $\alpha = 0$ e $\beta = 1$ temos a função densidade de probabilidade da distribuição de Cauchy padrão:

$$f(x) = \frac{1}{\pi\{1+x^2\}}, \quad -\infty < x < \infty,$$

Observe que $\frac{d}{dx} arctg(x) = \frac{1}{1+x^2}$ e portanto

$$\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1}{1+x^2} dx = \lim M \to \infty \int_{-M}^{M} \frac{1}{1+x^2} dx =$$

$$\frac{1}{\pi} \lim M \to \infty [arctg(M) - arctg(-M)] = \frac{1}{\pi}[\frac{\pi}{2} + \frac{\pi}{2}] = 1$$

com função de distribuição

$$F(x) = \frac{1}{\pi}\int_{-\infty}^{x} \frac{1}{1+x^2} dx = \frac{1}{2} + \frac{1}{\pi} arctg(x).$$

A distribuição de Cauchy é ilustrativa pois é exemplo de uma distribuição cujos momentos não existem, por exemplo

$$E[|X|] = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{|x|}{1+x^2} dx =$$

$$\frac{2}{\pi}\int_{0}^{\infty} \frac{x}{1+x^2} dx = \frac{\ln(1+x^2)}{\pi}|_0^{\infty} = \infty$$

Distribuição exponencial.

Como no caso discreto, uma variável aleatória contínua fica completamente caracterizada através de sua função de distribuição. Introduzimos a distribuição exponencial:

Definição 1.17. Uma variável aleatória T tem função de distribuição exponencial se, e somente se,

$$F_T(t) = \begin{cases} 0 & : \quad t < 0 \\ 1 - \exp[-\lambda t] & : \quad t \geq 0 \end{cases}.$$

$F_T(t) = 1 - \exp[-\lambda.t]$.

$F_T(t)$ é diferenciável e tem função densidade de probabilidade

$$f(t) = \begin{cases} 0 & : \quad t < 0 \\ \lambda \exp[-\lambda t] & : \quad t \geq 0 \end{cases}.$$

A função geradora de momentos de T é dada por

$$M_T(t) = E[\exp[t.T]] = \int_0^{\infty} \exp[tx]\lambda \exp[-\lambda.x] dx =$$

$$\int_0^\infty \lambda \exp[-x.(\lambda - t)]dx = \frac{\lambda}{\lambda - t}, t < \lambda.$$

com primeira e segunda derivadas iguais a $M_T^{(1)}(t) = \frac{\lambda}{(\lambda-t)^2}$ e $M_T^{(2)}(t) = \frac{2.\lambda.(\lambda-t)}{(\lambda-t)^4}$, respectivamente. No ponto zero temos

$$\mu = E[T] = M_T^{(1)}(0) = \frac{1}{\lambda}$$

e $E[T^2] = M_T^{(2)}(0) = \frac{2}{\lambda^2}$ de forma que

$$\sigma^2 = Var(T) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Observe que a média de T é o inverso de seu parâmetro.

Exemplo 1.18. Uma Cia. de Seguros tem observado que o custo dos sinistros de aeronaves de porte médio é modelado por uma variável aleatória com função de distribuição

$$F(t) = \begin{cases} 0 & : \quad t < 0 \\ 1 - \exp[-0,001t] & : \quad t \geq 0 \end{cases},$$

isto é, X tem distribuição exponencial. A probabilidade de um sinistro com custo maior do que 4.000 é

$$P(X > 4.000) = 0,001 \int_{4.000}^\infty \exp[-0,001x]dx = \exp[-4] = 0,018.$$

Se escolhemos casualmente 389 aeronaves seguradas, qual a probabilidade de que entre elas 10 reclamem sinistros com custos maiores do que 4.000? Se Y é definida como o número de aeronaves, dentre as 389, que reclamem sinistros com custos maiores do que 4.000, Y tem distribuição binomial com parâmetros $n = 389$ e $p = 0,018$ e $P(Y = 10) = \binom{389}{10}(0,018)^{10}(0,982)^{379}$, de cálculo difícil. Como $n = 389$ é grande e $p = 0,018$ é pequeno podemos aproximar a distribuiç o binomial pela distribuição de Poisson com média, a mesma de binomial, $\lambda = n.p = 389.0,018 = 7$. Assim, se N tem distribuição de Poisson com $\lambda = 7$,

$$P(Y = 10) = P(N = 10) = \frac{\exp[-7]7^{10}}{10!} = 0,07.$$

Introduzimos este capítulo exemplificando que a distribuição exponencial é uma extensão contínua da distribuição geométrica e como tal, tambem apresenta a propriedade de falta de memória. Uma variável aleatória T contínua e positiva que modela, por exemplo, o tempo de funcionamento de um componente eletrônico tem a propriedade de falta de memória se, e somente se,

$$P(T > t + s|T > s) = P(T > t),$$

onde $P(T > t) = 1 - P(T \leq t)$.

Teorema 1.19. *T é uma variável aleatória contínua com a propriedade de falta de memória se, e somente se, T tem distribuição exponencial.*

Prova *Se denotamos $g(t) = P(T > t)$ temos que $1 \leq g(t) \leq 1$ e que $g(t + s) = g(t)g(s)$.*

Para todo número natural n, temos

$$g(n) = g(1 + 1 + ... + 1) = g(1).g(1)....g(1) = g(1)^n.$$

Portanto $g(1) \neq 1$ pois, caso contrário, teríamos $0 = \lim_{n\uparrow\infty} g(n) = \lim_{n\uparrow\infty} g(1)^n = 1$, que é uma contradição.

Em adição

$$g(1) = g(\frac{n}{n}) = g(\frac{1}{n} + \frac{1}{n} + ... + \frac{1}{n}) = g(\frac{1}{n})^n,$$

o que implica $g(\frac{1}{n}) = g(1)^{\frac{1}{n}}$.

Portanto $g(1) \neq 0$ pois, caso contrário, teríamos $1 = \lim_{n\uparrow\infty} g(\frac{1}{n}) = \lim_{n\uparrow\infty} g(1)^{\frac{1}{n}} = 0$, que tambem é uma contradição.

Concluimos então que $0 < g(1) < 1$ e que existe um número real positivo λ, a imagem inversa de $g(1)$ através da função exponencial $f(x) = \exp[-x]$ tal que $g(1) = \exp[-\lambda]$

Pelo mesmo argumento, se n e m são números naturais

$$g(\frac{n}{m}) = g(\frac{1}{m} + \frac{1}{m} + ... + \frac{1}{m}) = g(\frac{1}{m})^n = g(1)^{\frac{n}{m}},$$

e concluimos que $g(x) = \exp[-\lambda x], \forall x \in \mathbf{Q}$, onde $\mathbf{Q}$ é o conjunto dos números racionais. Como $\mathbf{Q}$ é denso em $\Re$, temos

$$g(t) = P(T > t) = \exp[-\lambda t], \forall t \in \Re^+.$$

Consequentemente, a função de distribuição de T é

$$F(t) = \begin{cases} 0 & : \quad t < 0 \\ 1 - \exp[-\lambda t] & : \quad t \geq 0 \end{cases}.$$

que caracteriza completamente a distribuição exponencial. $\lambda > 0$ é o parâmetro da distribuição.

Vale o reverso, isto é, se T tem distribuição exponencial, temos

$$P(T > t{+}s) = \exp[-\lambda(t{+}s)] = \exp[-\lambda t].\exp[-\lambda s] = P(T > t).P(T > s)$$

e T tem falta de memória.

Exemplo 1.20. Uma industria fabrica lâmpadas especiais que ficam continuamente em operação. Caso a lâmpada dure menos do que 50 horas oferece a seus clientes a garantia de reposição. O tempo de vida útil dessas lâmpadas é modelado através da distribuição exponencial com parâmetro $\frac{1}{8.000}$. Portanto, a proporção de lâmpadas trocadas por garantia é

$$P(T \leq 50) = \int_0^\infty \frac{1}{8.000}\exp[-\frac{t}{8.000}]dt = 1 - \frac{1}{8.000} exp[-\frac{50}{8.000}] = 0,006.$$

Você acha razoável substituir uma lâmpada que já durou 5.000 horas? Como sabemos que a distribuição exponencial tem a propriedade da falta de memória, a resposta é não pois uma lâmpada usada é tão boa quanto uma nova. Analiticamente podemos escrever

$$P(T > t + 5.000 | T > 5.000) = \frac{P(T > t + 5.000, T > 5.000)}{P(T > 5.000)} =$$

$$\frac{\exp[-\frac{1}{8.000}(t + 5.000)]}{\exp[-\frac{1}{8.000}5.000]} = \exp[-\frac{1}{8.000}t] = P(T > t).$$

Distribuição de Pareto.

Definição 1.21. A função de distribuição de uma variável aleatória X, com função de distribuição de Pareto, $F(x)$ é definida por:

$$F(x) = 1 - (\frac{\theta}{x + \theta})^\alpha \quad x > 0,$$

Onde $\alpha > 0$ é o parâmetro da forma e $\theta > 0$ é o parâmetro de escala de sua função densidade de probabilidade

$$f(x) = \frac{\alpha\theta^\alpha}{(x + \theta)^\alpha}, \quad x > 0.$$

Portanto a esperança de X é

$$E[X] = \int_0^\infty x f(x)dx = \int_0^\infty (\int_0^x dy) f(x)dx = \int_0^\infty (\int_y^\infty f(x)dx)dy =$$

$$\int_0^\infty P(X > y)dy = \int_0^\infty \frac{\theta^\alpha}{(y + \theta)^\alpha} dy =$$

$$\theta^\alpha \int_\theta^\infty \frac{1}{y^\alpha} dy = \frac{\theta}{\alpha - 1}, \quad se \quad \alpha > 1.$$

Em geral vale

$$E[X^k] = \frac{\theta^k k!}{(\alpha - 1).(\alpha - 2)...(\alpha - k)}, \quad para \quad \alpha > k.$$

Concluimos que a variância da distribuição de Pareto é

$$\sigma^2 = Var(X) = E[X^2] - E[X]^2 = \frac{2.\theta^2}{(\alpha-1).(\alpha-2)} - (\frac{\theta}{\alpha-1})^2.$$

O p-ésimo percentil de X, x_p, é definido pela relação

$$F(x_p) = 1 - (\frac{\theta}{x_p+\theta})^\alpha = p$$

que tem solução

$$x_p = \theta.((1-p)^{-\frac{1}{\alpha}} - 1).$$

Exemplo 1.22. A perda aleatória, X, com seguro é modelada por uma distribuição de Pareto com média $\mu = 100.000$ e variância igual a $3.(10)^{10}$ unidades monetárias. Os parâmetros da distribuição podem ser calculados equacionando $\frac{\theta}{\alpha-1} = 100.000$ e $\frac{2.\theta^2}{(\alpha-1).(\alpha-2)} - (100.000)^2 = 3.(10)^{10}$ com solução $\alpha = 3$ e $\theta = 200.000$.

O atuário pode estimar a perda mediana calculando

$$x_{0,5} = 200.000((0,5)^{-\frac{1}{3}-1} = 52.000$$

e tambem se prevenir contra a pior perda definida por

$$E[X] + 3.\sigma = 100.000 + 3.\sqrt{3.(10)^{10}} = 619.615,24$$

com probabilidade $P(X > 619.616,24) = (\frac{200.000}{619.615,24+200.000})^3 = 0,0145.$

Se X é uma variável aleatória com distribuição de Pareto, a variável aleatória $Y = X + \theta$ assume valores $y > \theta$ e tem função de distribuição

$$P(Y \leq y) = P(X + \theta \leq y) = P(X \leq y - \theta) = 1 - (\frac{\theta}{y})^\alpha \quad y > \theta.$$

Neste caso, θ não é considerado um parâmetro, pois qualquer troca do valor de θ altera o dominio da distribuição definida em $[\theta, \infty)$ e y define a distribuição de Pareto uniparamétrica.

Exemplo 1.23. O tempo de vida de um sistema eletrônico depois de submetido a um processo de burn-in é modelado por uma distribuição de Pareto definida no intervalo $[1, \infty)$ com média $\mu = 5$.

Portanto o parâmetro α é definido pela equação

$$E[Y] = \int_1^\infty y \frac{\alpha}{y^{\alpha+1}} dy = -\frac{\alpha}{\alpha-1}|_1^\infty = \frac{\alpha}{\alpha-1},$$

resultando o valor $\alpha = 1,2$.

A probabilidade que o sistema sobreviva a 7 unidades de tempo é:

$$P(Y > 7) = \int_7^\infty \frac{1,2}{y^{2,2}} dy = 0,1.$$

Distribuição de Weibull.

Definição 1.24. Uma variável aleatória T com função de distribuição $F(t)$

$$F(t) = 1 - \exp[-(\lambda.t)^\tau] \quad t > 0$$

onde $\tau > 0$ é o parâmetro da forma e $\lambda > 0$ é o parâmetro de escala, tem função densidade de probabilidade

$$f(t) = \frac{dF(t)}{dt} = \lambda^\tau \tau t^{\tau-1} \exp[-(\lambda.t)^\tau], \quad t > 0$$

e é denominada distribuição de Weibull.

O momento de ordem k, de T pode ser calculado:

$$E[T^k] = \int_0^\infty t^k f(t)dt = \int_0^\infty k(\int_0^t s^{k-1}ds)f(t)dt =$$

$$k\int_0^\infty P(T > s)s^{k-1}ds = k\int_0^\infty \exp[-(\lambda.s)^\tau]s^{k-1}ds.$$

Substituindo $z = (\lambda.s)^\tau$, temos $s = \frac{z^{\frac{1}{\tau}}}{\lambda}$ e $ds = \frac{1}{\lambda.\tau}.z^{\frac{1}{\tau}-1}.dz$.
Portanto

$$E[T^k] = \frac{1}{\lambda^k}\frac{k}{\tau}\Gamma(\frac{k}{\tau}) = \frac{1}{\lambda^k}\Gamma(\frac{k}{\tau}+1)$$

onde a função Gama é definida como anteriormente: $\Gamma(t) = \int_0^\infty z^{t-1}\exp[-z], \quad t > 0$.

Consequentemente temos que

$$E[T] = \frac{1}{\lambda}\Gamma(\frac{1}{\tau}+1)$$

$$E[T^2] = \frac{1}{\lambda^2}\Gamma(\frac{2}{\tau}+1)$$

e

$$\sigma^2 = \frac{1}{\lambda^2.\tau}[2\Gamma(\frac{2}{\tau}) - \frac{1}{\tau}(\Gamma(\frac{1}{\tau}))^2].$$

Exemplo 1.25. Se T é uma variável aleatória representando uma perda, com função de distribuição

$$F(t) = 1 - \exp[-(0,02.t)^2] \quad t > 0,$$

calcule a perda média, a variância da perda e a probabilidade de que a perda não ultrapasse um desvio padrão da média.

A perda média é dada por

$$\mu = E[T] = \frac{1}{0,02}\Gamma(\frac{1}{2}+1) = 50.\frac{\sqrt{\pi}}{2} = 44,3,$$

$$E[T^2] = \frac{1}{(0,02)^2}\Gamma(\frac{2}{2}+1) = 2500$$

e

$$\sigma^2 = E[T^2] - (E[T])^2 = 2500 - (44,3)^2 = 537,5$$

com desvio padrão $\sigma = 23,2$.

A probabilidade de que T não ultrapasse $\mu + \sigma = 67,5$ é $F(67,5) = 1 - \exp[-(0,02.67,5)^2 = 0,84$.

Distribuição Normal.

Definição 1.26. Dizemos que uma variável aleatória X tem distribuição normal com parâmetros μ e σ^2, e denotamos $X \sim N(\mu, \sigma^2)$, se X tem função densidade de probabilidade

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp[-\frac{(x-\mu)^2}{2\sigma^2}], \;\; -\infty < x < \infty, \;\; \sigma > 0, \;\; -\infty < \mu < \infty.$$

A expressão de $f(x)$ parece complexa mas tem um gráfico suave em forma de um sino.

Figura 7.1 - Gráfico da curva normal

O parâmetro μ é de escala e o gráfico de normais com mesmo σ e diferentes μ são como segue:

Figura 7.2 - Gráficos de curvas normais

O parâmetro σ^2 é o da forma da densidade e quanto maior o σ a densidade é mais dispersa em torno de μ e existe chances maiores de encontrar valores , da variável, distantes de μ. S σ é pequeno, a densidade tem forma mais concentrada em torno de μ e a chance de observarmos valores próximos de μ aumenta.

Na realidade estamos definindo uma classe de distribuições normais pois a cada número real μ e número real positivo σ, temos uma distribuição normal. Em particular a distribuição normal com $\mu = 0$ e $\sigma^2 = 1$, denominada de distribuição normal padrão e denotada por $Z \sim N(0,1)$ tem função densidade de probabilidade

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp[-\frac{x^2}{2}], \quad -\infty < x < \infty.$$

Para ter uma noção da complexidade analítica de tal expressão provemos que

$$I = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp[-\frac{x^2}{2}] dx = 1.$$

Para proceder com a prova notemos que $I > 0$ e que $I = 1$ se e somente se, $I^2 = 1$. Contudo

$$I^2 = (\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp[-\frac{x^2}{2}] dx).(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp[-\frac{y^2}{2}] dy) =$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{2\pi} \exp[-\frac{1}{2}(x^2 + y^2)] dxdy.$$

Neste ponto consideramos uma transformação para coordenadas polares através das equações

$$x = \rho \sin\theta \quad e \quad y = \rho\cos\theta, \quad 0 \leq \theta \leq 2\pi, \quad 0 < \rho < \infty,$$

produzindo

$$I^2 = \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} \exp[-\frac{1}{2}\rho^2]\rho d\rho d\theta =$$

$$\frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} \exp[-y] dy d\theta = \frac{1}{2\pi}\int_0^{2\pi} d\theta = 1.$$

A função de distribuição dessa variável, $Z \sim N(0,1)$, é dada por

$$F(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp[-\frac{x^2}{2}]dx$$

mas o cálculo analítico para tal expressão é mais complexo do que os cálculos anteriores e devemos utilizar técnicas numéricas. Na prática, simplificamos com o uso das tabelas estatísticas, apresentadas no final do livro.

A tabela fornece a área sob a função densidade de probabilidade entre o valor zero 0 e um valor real z à sua direita. O valor z, até sua primeira casa decimal é encontrado na primeira coluna e sua segunda casa decimal na primeira linha da tabela. O cruzamento dessa linha e coluna, no interior da tabela, nos da a probabilidade $P(0 < Z \leq z)$. Procedendo dessa maneira verificamos, por exemplo:

Figura 7.3 - Cálculo de probabilidades

$$P(0 < Z \leq 1,64) = 0,45;$$
$$P(0 < Z \leq 1,96) = 0,475.$$

Desde que a função densidade de probabilidade da distribuição normal padrão é perfeitamente simétrica em relação a zero, concluimos que;

$$P(Z \leq -1,64) = P(Z \geq 1,64) = 0,5 - P(0 < Z \leq 1,64) = 0,05;$$

$$P(-1,96 \leq Z \leq 1,96) = 2.P(0 < Z \leq 1,96) = 0,95;$$

$$P(-1,64 \leq Z \leq 1.96) = P(-1,64 \leq Z \leq 0) + P(0 < Z \leq 1,96) =$$
$$P(0 < Z \leq 1,64) + P(0 < Z \leq 1,96) = 0,975.$$

Reversamente, podemos encontrar o valor de z, tal que a área (probabilidade) à sua esquerda, ou direita seja fixada. Consideremos encontrar z, tal que $P(Z \leq z) = 0,8$. Observe que este valor de z deve ser positivo, caso contrário, a área à sua esquerda seria menor ou igual a $0,5$. Como z é positivo podemos escrever

$$P(Z \leq z) = P(Z \leq 0) + P(0 < Z \leq z) = 0,5 + P(0 < Z \leq z) = 0,8,$$

que implica $P(0 < Z \leq z) = 0,3$. Observando o valor mais próximo de $0,3$ no corpo da tabela, percorremos sua linha e sua coluna, no sentido contrário do que vínhamos fazendo, No caso obtemos $z = 0,85$.

Para encontrar o valor de z com $P(Z \leq z) < 0,5$ podemos proceder com os mesmos argumentos, por exemplo:

Se $P(Z \leq z) = 0,05$, temos que z é negativo e que $P(Z \geq -z) = 0,05$. Portanto $P(0 < Z \leq -z) = 0,45$ e concluimos que $-z = 1,64$ e $z = -1,64$.

Conhecemos que a distribuição de uma variável aleatória é completamente determinada pela sua função geradora de momentos, $M_Z(t)$, quando esta função existir. A função geradora de momentos de Z é

$$M_Z(t) = E[\exp[tZ]] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp[tz] \exp[-\frac{z^2}{2}]dz =$$
$$\exp[\frac{t^2}{2}] \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp[-\frac{(z-t)^2}{2}]dz = \exp[\frac{t^2}{2}].$$

Podemos calcular a média e a variâcia de Z através da sua função geradora de momentos. A primeira derivada de $M_Z(t)$ é igual a $M_Z^{(1)}(t) = t.\exp[\frac{t^2}{2}]$ com $\mu = E[Z] = M_Z^{(1)}(0) = 0$.

A segunda derivada de $M_Z(t)$ é $M_Z^{(2)}(t) = \exp[\frac{t^2}{2}] + t^2.\exp[\frac{t^2}{2}]$ que no ponto zero vale $E[Z^2] = M_Z^{(2)}(0) = 1$. Portanto $\sigma^2 = var(Z) = E[Z^2] - \mu^2 = 1$.

Se consideramos a transformação linear da variável aleatória Z, $X = \sigma.Z + \mu$ onde μ é um número real e σ , um número real positivo, temos que a função geradora de momentos de X é

$$M_X(t) = E[\exp[tX] = E[\exp[t.(\sigma.Z + \mu)] = E[\exp[t.\sigma.Z]\exp[t.\mu]] =$$

$$\exp[t.\mu]M_Z(t.\sigma) = \exp[t.\mu + \frac{t^2\sigma^2}{2}].$$

Acontece que tal função geradora de momentos caracteriza completamente a distribuição normal com média μ e variãncia σ^2. Calculando as derivadas da função geradora de momentos no ponto zero, obtemos $E[X] = M_X^{(1)}(0) = \mu$ e $E[X^2] = M_X^{(2)}(0) = \sigma^2 + \mu^2$ e concluimos que os parâmetros da distribuição Normal, denotada por $X \sim N(\mu, \sigma^2)$ são sua média e variância.

Inversamente, podemos considerar a transformação reversa, $Z = \frac{X-\mu}{\sigma} = \frac{X}{\sigma} - \frac{\mu}{\sigma}$ e temos

$$M_Z(t) = E[\exp[t.Z]] = E[\exp[t.(\frac{X}{\sigma} - \frac{\mu}{\sigma})]] = \exp[-t.\frac{\mu}{\sigma})]].M_X(\frac{t}{\sigma}) = \exp[\frac{t^2}{2}].$$

Ccncluimos que existe uma equivalência entre a normal de média μ e variância σ^2, $N(\mu, \sigma^2)$, e a normal padrão, $N(0, 1)$, através das citadas transformações e utilizaremos este resultado para cálculos envolvendo a função de distribuição de $X \sim N(\mu, \sigma^2)$ através da função de distribuição da normal padrão $Z \sim N(0, 1)$.

Assim, se $X \sim N(\mu, \sigma^2)$, a sua função de distribuição pode ser calculada como

$$F_X(x) = P(X \leq x) = P(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}) = P(Z \leq \frac{x - \mu}{\sigma}) = F_Z(\frac{x - \mu}{\sigma}).$$

Exemplo 1.27. O custo dos sinistros de certo tipo de apólice tem distribuição normal com média de $R\$1.800,00$ e desvio padrão $R\$400,00$. A probabilidade de que um sinistro escolhido aleatóriamente custe mais do que $R\$1.500,00$ é

$$P(X > 1500) = P(\frac{X - \mu}{\sigma} > \frac{1500 - 1800}{400}) = P(Z > -0,75) =$$

$$P(Z \leq 0,75) = 0,5 + P(0 < Z \leq 0,75) = 0,5 + 0,27337 = 0,77.$$

A probabilidade de que o custo esteja entre $R\$1.500,00$ e $R\$2.000,00$ é

$$P(1500 < X \leq 2000) = P(\frac{1500-1800}{400} < \frac{X-\mu}{\sigma} \leq \frac{2000-1800}{400}) =$$
$$P(-0,75 < Z \leq 0,5) = P(-0,75 < Z \leq 0) + P(0 < Z \leq 0,5) =$$
$$P(0 < Z \leq 0,75) + 0,19146 = 0,27337 + 0,19146 = 0,46.$$

A probabilidade de que o custo esteja a $1,96$ desvio padrão de sua média é

$$P(|X-1800| \leq 1,96.400) = P(-1,96.400 \leq X - 1800 \leq 1,96.400) =$$
$$P(-1,96 \leq \frac{X-1800}{400} \leq 1,96) = P(-1,96 \leq Z \leq 1,96) = 0,95.$$

Caso a Cia de Seguros deseje estabelecer uma franquia de forma que 10% das ocorrências não utilizem o seguro, devemos procurar o custo x tal que $P(X \leq x) = 0,1)$.

Padronizando X, a equação é equivalente a

$$P(\frac{X-\mu}{\sigma} \leq \frac{x-1800}{400}) = 0,1 \leftrightarrow P(Z \leq \frac{x-1800}{400}) = 0,1 \leftrightarrow$$
$$P(Z \geq \frac{x-1800}{400}) = 0,1 \leftrightarrow P(0 < Z \leq \frac{1800-x}{400}) =$$
$$0,4 \leftrightarrow \frac{1800-x}{400} = 1,64 \leftrightarrow x = 1.144,$$

isto é, a franquia deve ser de $R\$1.144,00$.

Observação 1.28. **Aproximação da distribuição binomial pela distribuição normal.**

Consideramos no exemplo 1.8 do Capítulo anterior o cálculo da probabilidade de que entre 10.000 segurados, 15 solicitem benefícios devido à ocorrências de sinistros. A probabilidade da ocorrência de tais sinistros foi estipulada como $p = 0,001$. Se Y é a variável aleatória definida pelo número de segurados que reclamam os benefícios, temos que Y tem distribuição binomial de parâmetros 10.000 e $0,001$ ($Y \sim B(10.000, 0,001)$).

A probabilidade do evento $\{Y = 15\}$ é

$$P(Y = 15) = \binom{10.000}{15} 0,001^{15} 0,999^{9985},$$

de cálculo complicado. Naquele exemplo aproximamos tal probabilidade pela probabilidade da distribuição de Poisson com mesma média da distribuição binomial. Neste Capítulo consideramos a aproximação da distribuição binomial pela distribuição normal com mesma média e variância da binomial.

Para um melhor entendimento consideremos a distribuição binomial de parâmetros $n = 10$ e $p = 0,5$. A função de probabilidade da variável aleatória $Y \sim B(10; 0,5)$ é

y	0	1	2	3	4	5	6	7	8	9	10
$P(Y = y)$	$0,001$	$0,01$	$0,044$	$0,117$	$0,205$	$0,246$	$0,205$	$0,117$	$0,044$	$0,01$	$0,001$

com representação gráfica:

Figura 7.4- Aproximação da binomial pela normal

Uma primeira aproximação gráfica de uma variável discreta por uma variável contínua é através do histograma, o gráfico de retângulos contíguos cujas áreas somam 1. Se para cada valor da variável discreta construirmos um retângulo de base 1 e altura igual à sua probabilidade teremos a soma dessas áreas igual a 1. A base do retângulo correspondente ao valor x_i, da variável, é definida por $x_i - 0,5$ e $x_i + 0,5$, como na figura acima.

Seja X a variável aleatória que corresponde ao histograma. No caso em que n é grande a variável X converge para uma variável normal com média $\mu = np$ e $\sigma^2 = n.p.(1 - p)$, a média e variância da distribuição $B(n, p)$.

No exemplo em que $n = 10$ e $p = 0,5$ temos $P(Y = 5) = 0,246$, $P(Y \leq 3) = 0,172$ e $P(2 \leq Y < 6) = 0,612$.

Se $X \sim N(5; 2,5)$,

$$P(4,5 \leq X \leq 5,5) = P(\frac{4,5-5}{1,58} \leq \frac{X-5}{1,58} \leq \frac{5,5-5}{1,58}) =$$
$$P(-0,32 \leq Z \leq 0,32) = 2.P(0 \leq Z \leq 0,32) = 2.0,12552 = 0,251,$$

uma aproximação por menos de 5 milésimos.

$$P(X \leq 3,5) = P(\frac{X-5}{1,58} \leq \frac{3,5-5}{1,58}) = P(Z \leq -0,95) =$$
$$P(Z > 0,95) = 0,5 - P(0 \leq Z \leq 0,95) = 0,5 - 0,32894 = 0,171.$$

$$P(1,5 \leq X \leq 5,5) = P(\frac{1,5-5}{1,58} \leq \frac{X-5}{1,58} \leq \frac{5,5-5}{1,58}) =$$
$$P(-2,21 \leq Z \leq 0,32) = P(0 \leq Z \leq 2,21)+$$
$$P(0 \leq Z \leq 0,32) = 0,48645 + 0,12552 = 0,61.$$

A aproximação colocada tem justificativa rigorosa quando demonstramos o Teorema do Limite Central que aproxima soma de variáveis aleatória independentes e identicamente distribuidas por uma distribuição normal. Observe que uma variável binomial pode ser interpretada como soma de Bernoulli que são independentes e identicamente distribuidas.

Portanto, considerando o exemplo 8 do Capítulo anterior, poderíamos aproximar $Y \sim B(10.000; 0,001)$ pela distribuição normal $X \sim N(10; 9,99)$ e

$$P(Y = 15) \approx P(14,5 \leq X \leq 15,5) = P(\frac{14,5-10}{3,16} \leq \frac{X-10}{3,16} \leq \frac{15,5-10}{3,16}) =$$
$$P(1,4 \leq Z \leq 1,74) = P(Z \leq 1,74) - P(Z \leq 1,4) = 0,45907 - 0,41924 = 0,0398.$$

O resultado da aproximação pela distribuição de Poisson foi $0,0347$.

Teorema 1.29. *Se Z é uma variável aleatória contínua com distribuição normal com média $\mu = 0$ e variância $\sigma^2 = 1$, a variável aleatória Y, resultado da transformação $Y = Z^2$, tem função densidade de probabilidade*

$$f(y) = \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{y}} \exp[-\frac{y}{2}], \quad y > 0,$$

denominada de distribuição qui-quadrado com um grau de liberdade.

Prova *Evidentemente, os valores de Y são positivos e a função de distribuição de Y é dada por*

$$P(Y \leq y) = P(Z^2 \leq y) = P(-\sqrt{y} \leq Z \leq \sqrt{y}) =$$

$$\int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} \exp[-\frac{x^2}{2}]dx = 2.\int_{0}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} \exp[-\frac{x^2}{2}]dx.$$

Portanto a função densidade de probabilidade de Y, $f(y) = \frac{dF(y)}{dy}$ *é*

$$f(y) = \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{y}} \exp[-\frac{y}{2}], \quad y > 0,$$

Distribuição qui-quadrado.

Definição 1.30. A função de distribuição de uma variável aleatória Y com função densidade de probabilidade

$$f(y) = \frac{1}{\Gamma(\frac{k}{2})} (\frac{1}{2})^{\frac{k}{2}} y^{\frac{k}{2}-1} \exp[-\frac{y}{2}], \quad y > 0,$$

é denominada de distribuição qui-quadrado com k graus de liberdade, tem média $\mu = k$ e variância $\sigma^2 = 2.k$.

A função Gama, denotada por Γ que aparece na definição acima é definida por

$$\Gamma(t) = \int_0^{\infty} x^{t-1} \exp[-x]dx, \quad t > 0.$$

Por integração por partes pode-se provar que $\Gamma(t+1) = t\Gamma(t)$. Se $t = n$, um número natural, temos $\Gamma(n+1) = n!$. Em particular $\Gamma(n+\frac{1}{2}) = \frac{1.3.5....(2n-1)}{2^n}\sqrt{2\pi}$ e $\Gamma(\frac{1}{2}) = 2.\Gamma(\frac{3}{2}) = \sqrt{\pi}$.

Distribuição gama.

Definição 1.31. A função de distribuição de uma variável aleatória X com função densidade de probabilidade

$$f(x) = \frac{1}{\Gamma(\alpha)} \lambda^{\alpha} x^{\alpha-1} \exp[-\alpha y], \quad x > 0,$$

é denominada de distribuição gama com parâmetros α e λ, os quais são números reais positivos.

A distribuição gama é denotada por gama(α, λ) e tem média $E[X] = \frac{\alpha}{\lambda}$ e variância $\sigma^2 = \frac{\alpha}{\lambda^2}$.

Distribuição Beta.

Uma variável aleatória caracterizada pela função densidade de probabilidade

$$f(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}, \quad 0 < x < 1,$$

onde $a > 0$ e $b > 0$ tem função de distribuição denominada Beta de parâmetros a e b. A função é definida por

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

e portanto $\int_0^1 f(x)dx = 1$.

Teorema 1.32. *Se X é uma variável aleatória com distribuição Beta, de parâmetros a e b, então*

$$E[X^k] = \frac{\Gamma(k+a)\Gamma(a+b)}{\Gamma(a)\Gamma(k+a+b)}.$$

Prova

$$E[X^k] = \frac{1}{B(a,b)} \int_0^1 x^{k+a-1}(1-x)^{b-1}dx = \frac{B(k+a,b)}{B(a,b)} =$$

$$\frac{\Gamma(k+a)\Gamma(b)}{\Gamma(k+a+b)} . \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} =$$

$$\frac{\Gamma(k+a)\Gamma(a+b)}{\Gamma(a)\Gamma(k+a+b)}$$

e concluimos a prova.

Usando o teorema temos

$$E[X] = \frac{\Gamma(1+a)\Gamma(a+b)}{\Gamma(a)\Gamma(1+a+b)} = \frac{a}{a+b},$$

$$E[X^2] = \frac{\Gamma(2+a)\Gamma(a+b)}{\Gamma(a)\Gamma(2+a+b)} = \frac{(a+1)a}{(a+b+1)(a+b)},$$

e

$$Var(X) = \frac{(a+1)a}{(a+b+1)(a+b)} - (\frac{a}{a+b})^2 = \frac{ab}{(a+b+1)(a+b)^2}.$$

Um resultado interessante relaciona a distribuição binomial e a distribuição beta:

$$\Sigma_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j} = \frac{1}{B(k, n-k+1)} \int_0^p x^{k-1}(1-x)^{n-k}dx.$$

Prova

Por integração por partes, se

$$v = (1-x)^{n-k} \Rightarrow dv = (n-k)(1-x)^{n-k-1}(-1)dx$$

e se

$$du = x^{k-1}dx \Rightarrow u = \frac{x^k}{k}$$

, temos:

$$\frac{1}{B(k,n-k+1)} \int_0^p x^{k-1}(1-x)^{n-k}dx =$$

$$\frac{1}{B(k,n-k+1)}\{\frac{p^k}{k}(1-p)^{n-k} + (\frac{n-k}{k}) \int_0^p x^k(1-x)^{n-k-1}dx\} =$$

Observando que $\frac{1}{B(k,n-k+1)} = \binom{n}{k}k$ e repetindo o processo de integração por partes temos:

$$\binom{n}{k}p^k(1-p)^{n-k} + \binom{n}{k+1}[p^{k+1}(1-p)^{n-k-1}+$$

$$\frac{n!}{(k+1)!(n-k-2)!}\int_0^p x^{k+1}(1-x)^{n-k-2}dx =$$

$$... = \Sigma_{j=k}^n \binom{n}{j} p^j(1-p)^{n-j}.$$

Observação 1.33. **Transformações de variáveis aleatórias**

O Teorema 6.28 pode ser colocado como um caso particular do Teorema que segue:

Teorema 1.34. *Suponha que X é uma variável aleatória do tipo contínuo com função de densidade de probabilidade $f_X(x)$, com domínio $\Re_X = \{x : f_X(x) > 0\}$. Assuma que:*

a) $y = g(x)$ define uma transformação um a um e sobrejetora de $\Re_X$ em $\Re_Y = \{y : g(y) > 0\}$.

b) A derivada de $x = g^{-1}(y)$, com respeito a y é contínua em $\Re_Y$.

Então $Y = g(X)$ é uma variável aleatória do tipo contínuo com função de densidade de probabilidade

$$f_Y(y) = |\frac{dg^{-1}(y)}{dy}|.f_X(g^{-1}(y)), \quad y \in \Re_Y.$$

Prova

Se $g(x)$ é crescente, $g^{-1}(y)$ é crescente. Então

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y))$$

e portanto

$$f_Y(y) = \frac{dF_X(g^{-1}(y))}{dy} = |\frac{dg^{-1}(y)}{dy}|.f_X(g^{-1}(y)).$$

No caso em que $g(x)$ é decrescente, $g^{-1}(y)$ é decrescente e

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \geq g^{-1}(y)) = 1 - F_X(g^{-1}(y))$$

e

$$f_Y(y) = \frac{dF_X(g^{-1}(y))}{dy} = |\frac{dg^{-1}(y)}{dy}|.f_X(g^{-1}(y)).$$

A restrição de que $y = g(x)$ seja bijetiva no domínio $\Re_X$ é restritiva. Se podemos particionar $\Re_X$ em $\Re_X^1, \Re_X^2, ..., \Re_X^m$ de maneira que em cada $\Re_X i$, $g(x)$ é bijetora, podemos aplicar o Teorema em cada $\Re_X i$ e concluir que

$$f_Y(y) = \sum_{i=1}^{m} |\frac{dg^{i^{-1}}(y)}{dy}|.f_X(g^{i^{-1}}(y)), \quad y \in \Re_Y.$$

Teorema 1.35. *Se X é uma variável aleatória contínua com distribuição normal com média μ e variância σ^2, a variável aleatória Y, resultado da transformação $Y = \exp[X]$, tem função densidade de probabilidade*

$$f(y) = \frac{1}{y.\sigma.\sqrt{2\pi}} \exp[-\frac{(-\mu)^2}{2\sigma^2}], \quad y > 0.$$

denominada de distribuição log-normal.

Prova *É evidente que os valores y que Y assume são positivos. A função de distribuição de Y é*

$$F(y) = P(Y \leq y) = P(\exp[X] \leq y) = P(X \leq \ln y) =$$

$$\int_{-\infty}^{\ln y} \frac{1}{\sigma\sqrt{2\pi}} \exp[-\frac{(x-\mu)^2}{2\sigma^2}]dx.$$

Portanto, a função densidade de probabilidade de Y é

$$f(y) = \frac{dF(y)}{dy} = \frac{1}{y.\sigma.\sqrt{2\pi}} \exp[-\frac{(-\mu)^2}{2\sigma^2}].$$

Definição 1.36. A função de distribuição de uma variável aleatória Y com função densidade de probabilidade

$$f(y) = \frac{1}{y.\sigma.\sqrt{2\pi}} \exp[-\frac{(-\mu)^2}{2\sigma^2}], \quad y > 0,$$

é denominada de distribuição Lognormal com parâmetros μ e σ^2.

Recordemos que a função geradora de momentos da variável aleatória $X \sim N(\mu, \sigma^2)$ é

$$M_X(t) = E[\exp[t.X]] = \exp[t.\mu + \frac{\sigma^2 t^2}{2}].$$

Portanto

$$E[Y] = E[\exp[X]] = M_X(1) = \exp[\mu + \frac{\sigma^2}{2}],$$

$$E[Y^2] = E[\exp[2.X]] = \exp[2.\mu + 2.\sigma^2]$$

e

$$\sigma_Y^2 = \exp[2.\mu + 2.\sigma^2] - \exp[2.\mu + \sigma^2].$$

Exemplo 1.37. A taxa de crescimento de uma população é uma variável aleatória Y com distribuição normal com média $0,03$ e variância 0,0001. O crescimento da população, com tamanho inicial de 100.000 indivíduos no período de um ano é modelado por $Y = 100.000 \exp[X]$. Reconhecemos que a variável $\frac{Y}{100.000}$ tem distribuição lognormal com os mesmos parâmetros de X e portanto, no período de um ano esperamos uma população de

$$E[Y] = 100.000.E[\exp[X]] = 100.000 \exp[0,03 + \frac{0,0001}{2}] = 103.051.$$

E-mail address: bueno@ime.usp.br

Departamento de Estatística, Instituto de Matemática e Estatística, Universidade de São Paulo, Caixa Postal 66281, CEP 05311-970, São Paulo, Brazil